

# Speech Processing 2012/13

1st Test

March 25th 2013

Please identify this form with your name and student number in the reserved space. The answers to multiple-choice questions will only be accepted if inserted in the appropriate place. Wrong answers will be penalized. The phonetic symbols should use the SAMPA alphabet (Lisbon accent).

1. Classify as True (T) or False (F)
  - (a) Narrowband spectrograms allow the discrimination of individual harmonics.
  - (b) The part of the basilar membrane which responds maximally to low frequency sounds is the stiff base.
  - (c) The ear is most sensitive to frequencies below 500 Hz.
  - (d) Voiced plosives have a lower VOT than unvoiced ones.
  - (e) Vocal fry is a pathological mode of voiced excitation.
2. Name three phenomena that may justify considering context-dependent allophones.
3.
  - (a) What is the name of the recursive method which is typically used for computing the solution of the autocorrelation method of linear prediction?
  - (b) Which are the most efficient equivalent representations of linear prediction coefficients in terms of spectral distortion?
4. Write the expression that relates the real and complex cepstra of a signal.
5. Which time-domain parameters may be used to decide whether a segment is voiced or unvoiced?
6. Figure 1 shows waveforms and spectrograms of two syllables both ending in the same vowel. The initial sounds have the same manner and place of articulation (palatal) in the two syllables. Which are their corresponding SAMPA symbols (left to right)?

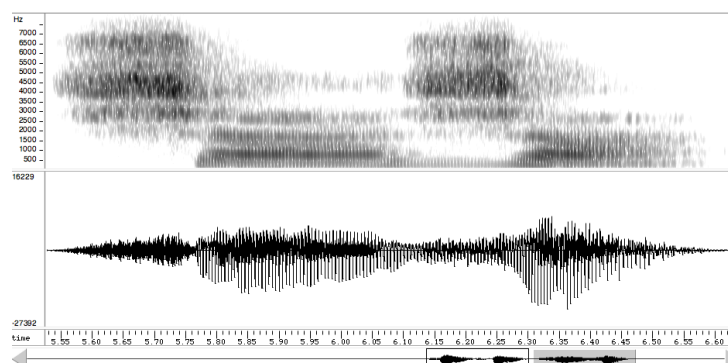


Figure 1.

7. Compute the length of a uniform lossless tube that has a third formant of the neutral vowel at 2700 Hz, indicating all steps.

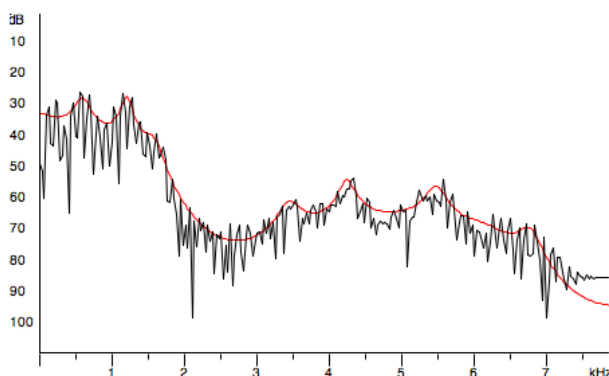
8. Consider the following sentence in European Portuguese:

*Amanhã as greves de transportes afectarão milhões de utentes em todo o país.*

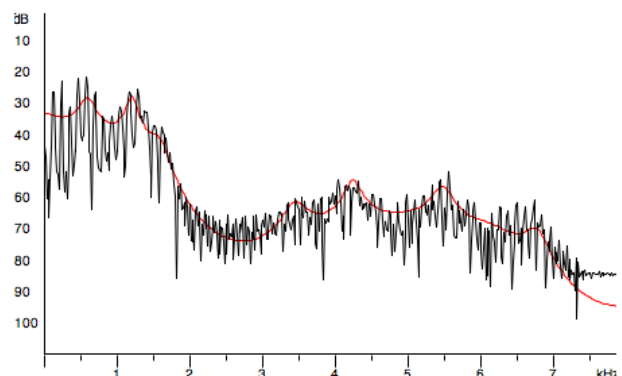
If you are not a native Portuguese speaker, you may use the Google translation in English, and indicate which SAMPA alphabet you are using (EP or EN):

*Tomorrow the transport strikes will affect millions of users across the country*

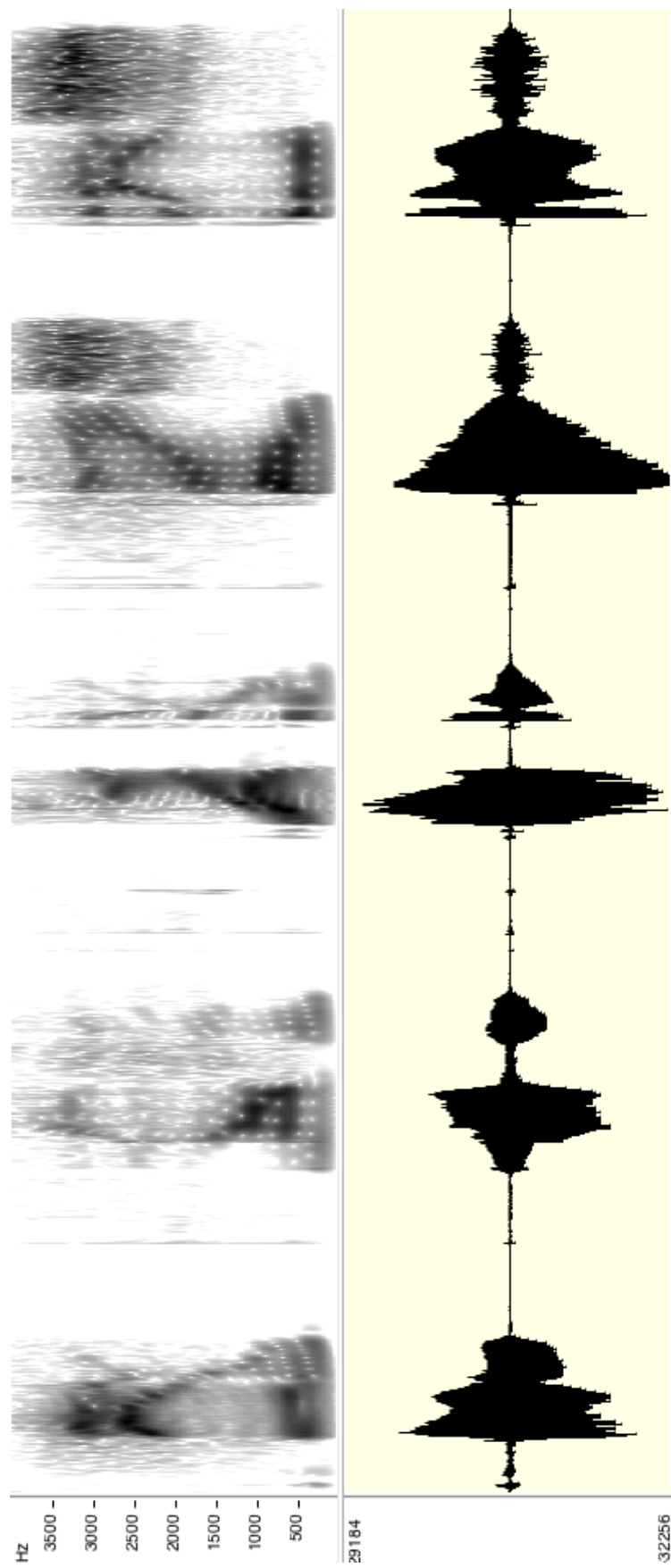
- (a) Write the broad phonetic transcription.
  - (b) Which of the following classes is less represented in this sentence (vowels / fricatives / plosives / liquids / laterals / nasal consonants)?
9. Identify the 10-digit sequence by inspecting the spectrograms and waveforms in pages 3 and 4. There are no repeated digits. The recordings correspond to telephone speech.
10. A sentence has been recorded with 8kHz, 16 bits per sample, duration 3.5s. A pitch detection method, with windows 20 ms long and interval 10 ms, computed values for all frames except for 4 boundary frames, yielding around 20% of unvoiced frames. A linear prediction analysis of order  $p$  was performed, and the corresponding residual was computed. (Please indicate all the steps you computed in answering the following questions).
- (a) What is the total size of the recorded file in bytes?
  - (b) How many non-zero pitch values were computed?
  - (c) What value of  $p$  would you suggest?
  - (d) Does the residual signal show periodicity in voiced frames?
11. Figures 2 present the short-time Fourier transform (magnitude) of a vowel segment, including its LPC spectral envelope.
- (a) Indicate approximate values for F0, F1 and F2 (Hz).
  - (b) To which of the 3 vowels of the extremes of the vowel triangle does it correspond?
  - (c) May it correspond to a male voice?
  - (d) May the spectral envelope be obtained with an LPC analysis of order 8?
  - (e) Does it correspond to wideband speech?
  - (f) Which of the figures *a* or *b* corresponds to the largest FFT analysis window?

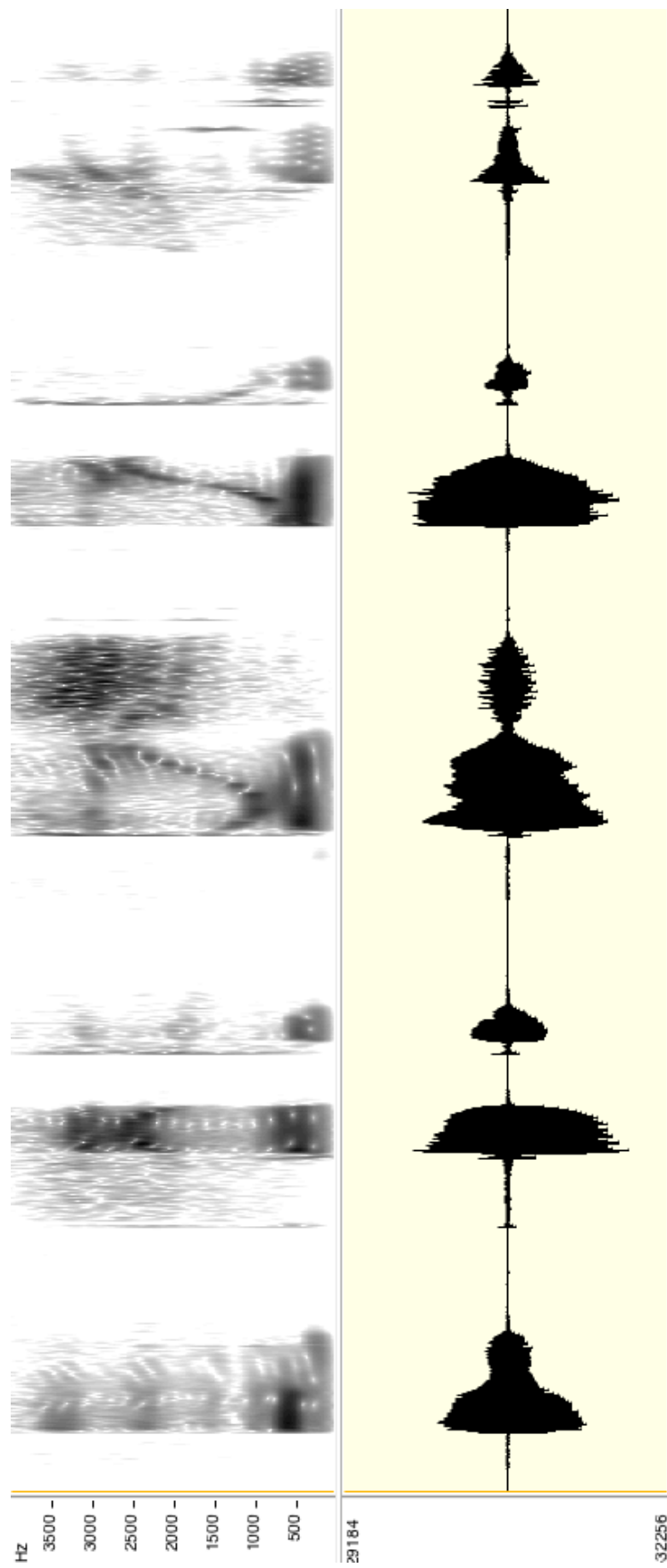


2a



2b





**Test 1 - Answers**

Name:	
Number:	
Group Number:	

1. (1.0 val.) Indicate T or F:

a	b	c	d	e

2 (1.5 val.)

a	
b	
c	

3 (1.4 val.)

a	
b	

4 to 7. (1.2/1.5/1.8/1.5)

4	
5	
6	
7	

8. (2.4 val.)

a	
b	

9. (3 val.)

--	--	--	--	--	--	--	--	--	--	--

10. (1.7 val.)

a	
b	
c	
d	

11. (3.0 val.) Complete:

F0	F1	F2	Vow	c	d	e	f